M3H/M4H/M5H HISTORY OF MATHEMATICS: EXAMINATION SOLUTIONS, 2016

Q1. *The Platonic solids*

How many solid figures are there whose faces are congruent regular polygons?

Triangular faces. How many faces (equilateral triangles) meet at a vertex? If 3, we have a *tetrahedron*; if 4, an *octahedron*; if 5, an *icosahedron* (2 or less would not give a solid; 6 gives instead a triangular tessellation of the plane, again not a solid; there is no room for 7 or more).

Square faces: 3 gives a cube (4: square tessellation of the plane – 'OS grid').

Pentagonal faces. 3 gives a dodecahedron (4 is impossible, as the angle at the vertex of pentagon is $3\pi/5$ and $4 \times 3\pi/5 > 2\pi$).

Hexagonal faces give a honeycomb tessellation of the plane, and not a solid. So the list above is exhaustive. **[3]**

These 5 'regular polyhedra' are called the *Platonic solids*, after Plato's use of them in the Dialogue of Timaeus. Apparently, according to a scholium to Book XIII of Euclid's Elements, the tetrahedron, cube and dodecahedron were known to the Pythagoreans, while Theaetetus (d. 369 BC) found the octahedron and icosahedron (he probably also proved the theorem above: the list is exhaustive). This is perhaps surprising: it is the dodecahedron and icosahedron that seem the hardest, while the octahedron is merely a pyramid (!) reflected in its base-plane. But archaeological support exists: an Etruscan dodecahedron made c. 500 BC of stone was found in 1885 on Monte Loffa near Padua, and there are Celtic examples. Perhaps also the idea of reflection is not as self-evident as it may seem today.

Unfortunately, Plato saw mystical significance here (earth: cube; air: octahedron; fire: tetrahedron; water: icosahedron; universe: dodecahedron) – a throwback to the superstition of the Pythagoreans and the Babylonians. **[3]**

Natural occurrence

Crystals: Tetrahedron: sodium sulphantimoniate; cube: common salt; octahedron: chrome alum **[1]**

Living forms: skeletons of radiolaria (microscopic sea creatures):
Circogonia: icosahedra; Circorrhegma: dodecahedra. **[1]**

Duality.

From the point of view of Projective Geometry: the tetrahedron is self-dual; the cube and octahedron are dual; the dodecahedron and icosahedron are dual. [Seen – lectures] **[2]**

Q2. *The Fundamental Theorem of Arithmetic, FTA.*
Theorem (FTA). Every integer $n \geq 2$ can be written uniquely (to within order) as a product of prime factors. **[1]**
Proof. Existence. Induction. True for $n = 2$. Assume true for every integer $< n$. If n is not prime (i.e. is composite), it has a non-trivial divisor d $(1 < d < n)$. So $n = cd$ $(1 < c < n)$. So each of c, d is a product of primes, by the inductive hypothesis. So n is too, completing the induction. **[1]**
Uniqueness. Induction. True for $n = 2$. Assume true for every integer $< n$. If n is prime, the result holds, so assume n is composite. If it has two factorisations $n = p_1 \dots p_r = q_1 \dots q_s,$
to show $r = s$ and each p is some q. As p_1 is prime and divides the product $n = q_1 \dots q_s$, it must divide at least one factor (w.l.o.g., q_1): $p_1 | q_1$. Then $p_1 = q_1$ as both p_1, q_1 are prime. Cancel p_1:

$$n/p_1 = p_2 \dots p_r = q_2 \dots q_s.$$

As n is composite, $1 < n/p_1 < n$. Then the inductive hypothesis tells us that the two factorisations above of n/p_1 agree to within order: $r = s$, and $p_2, \dots, p_r$ are $q_2, \dots, q_r$ in some order, as required. // **[1]**
Historical Note.

We owe Mathematics as a subject to the ancient Greeks. Of the 13 books of Euclid's Elements (EUCLID of Alexandria, c. 300 BC), three (Books VI, IX and X) are on Number Theory. From the ordering of the material in Euclid, it is clear that the Greeks knew that they did not have a proper theory of irrationals (i.e. reals). Although they did not state FTA, it had been assumed that they "knew it really", but did not state it explicitly. This view is contradicted by Salomon BOCHNER (1899-1982) (*Collected Papers*, Vol. 4, AMS, 1992). According to Bochner, the Greeks did *not* know FTA, nor have a notational system adequate even to state it! **[3]**

L. E. DICKSON (1874-1954) (*History of the Theory of Numbers* Vols 1-3, 1919-23) does not address the question of the Greeks and FTA! **[1]**

The first clear statement and proof of FTA is in Gauss' thesis (C. F. GAUSS (1777-1855); *Disquisitiones Arithmeticae*, 1798, publ. 1801). So we must attribute the result to Gauss. This is a wonderful example of the excellent mathematical *taste* of the young man who went on to be the greatest mathematician of all time. **[3]**
[Seen in less detail in lectures. The above is actually taken from my Analytic Number Theory notes, Lecture 1.]

Q3. *The Fundamental Theorem of Algebra.*
Carl Friedrich Gauss (1777-1855)

Gauss' doctoral thesis (in Latin: 'A new proof that every polynomial of one variable can be factored into real factors of the first or second degree') was published in 1799.

Despite its name, this result is a theorem of *analysis*, not of *algebra*. Its proof was less rigorous than Gauss' usual standard: he assumed properties of continuous functions later proved by Bolzano. **[3]**

Augustin-Louis Cauchy (1789-1857), Professor at the Ecole Polytéchnique and later the Sorbonne.

Cauchy's *Cours d'analyse* (Ecole Polytéchnique, 1821) contains a proof of the Fundamental Theorem of Algebra: every complex polynomial of degree n has n complex roots (counted according to multiplicity). **[3]**

The modern proof uses *Liouville's theorem* (Joseph Liouville (1809-1882), lectures in 1847 – actually published by Cauchy in 1844): an entire (i.e. holomorphic throughout the complex plane $\mathbb{C}$) bounded function is constant. Thus it took over twenty years before the full power of Cauchy's new subject of Complex Analysis was properly brought to bear on the Fundamental Theorem of Algebra – incidentally, revealing in so doing that the result, being a theorem in Analysis, is a misnomer. **[2]**

N. H. Abel (1802-29): Insolubility of the quintic (1829).

Although the quintic has five roots, by the Fundamental Theorem of Algebra above, Abel showed that – in contrast to polynomials of degree up to four, which can be solved, as Cardano showed – quintics are *not soluble by radicals*: there can be *no* formula/algorithm/method for expressing the roots in terms of the coefficients. Similarly for polynomials of higher degree. So there is a fundamental split: polynomial equations are soluble by radicals for degree up to 4, but not for degree 5 or higher. **[2]**

[Seen – lectures]

Q4. *The mathematics of perspective.*

Perspective in the Ancient World **[3]**

We know from the writings of classical authors, such as the Roman author Vitruvius, and from surviving wall- and vase-paintings, that some elements of perspective were known to the ancient Greeks. Agatharchos used it for stage sets in the late 5th C. BC. Theoretical studies of perspective were made by Anaxagoras (above) and Democritus (above).

W: Systematic attempts to evolve a system of perspective are usually considered to have begun around the fifth century BC. in the art of Ancient Greece, as part of a developing interest in illusionism allied to theatrical scenery and detailed within Aristotle's Poetics as 'skenographia'.

Perspective in the Renaissance. **[4]**

As we have seen, perspective was known (at least in part) in the ancient world, but was then lost.

Filippo Brunelleschi (1377-1446) discovered the main principle of perspective – the use of vanishing points – and convinced his fellow-artists of this in a famous experiament of 1420 involving the chapel outside Florence Cathedral.

Leon Battista Alberti (1404-72), *Della pictura* (1435, printed 1511) gave the first written account of perspective.

Piero della Francesca (1410-92), *De prospectivo pingendi* (c. 1478). In his book, and in his painting, Piero della Francesca did much to popularise perspective, which spread throughout the Western art world.

Leonardo da Vinci (1452-1519); *Trattato della pittura*. Leonardo is usually regarded as the personification of Renaissance genius. He was a prolific inventor, an artist who wrote on perspective, and a mathematician.

Albrecht Dürer (1471-1528) of Nuremburg; *Investigations of the measurement with circles and straight lines of plane and solid figures* (1525-1538, German and Latin). Like Leonardo, Dürer was both a mathematician and an artist. He adopted perspective after visiting Italy.

Girard Desargues (1591-1661) and *Projective Geometry.* **[3]**

Desargues' first important book was *La Perspective* (1636). This led him on to his introduction of projective geometry in his *Brouillon projet (d'une atteinte aux evenements des rencontres d'une cone avec un plan)* (1639) – *Rough draft (of an attempt to deal with the outcome of a meeting of a cone with a plane)*. Projective geometry is the mathematics of perspective (vanishing point = 'point at infinity').

[Seen – lectures]

Q5. *The mathematics of the rainbow.*

The rainbow is formed when the sun is behind the observer, and rain is falling ahead of the observer. One needs two physical principles:
(a) The law of reflection: when light is reflected at a mirror, the angle of incidence = angle of reflection. (b) Snel's law of refraction: "mu sin theta = constant", where mu is the refractive index (higher for water than for air, as light travels more slowly in water than in air), and theta is the angle of incidence): the Dutch scientist Willebrord Snel (1581-1625), 1618 and 1621. [**2**]

The rainbow has a definite 'size'. This is the angle between the line L_1 from the Sun through the observer and any line L_2 from the observer to the arc of the rainbow (any such line gives the same angle, as this arc is circular).

The rainbow is produced when light from the Sun is
(i) *refracted* when it enters falling raindrops in front of the observer (and is bent *towards* the normal);
(ii) *reflected* at the back of the raindrop (which acts as a mirror);
(iii) *refracted* again when it exits the raindrops.

A visible effect – the rainbow – is obtained when the angle of deviation has an *extremum* (minimum). For here, many rays will emerge parallel. [**2**]

Light of different colours (wavelengths) have different refractive indices, so are separated by the two refractions, and one sees the colours of the rainbow. The red end of the visible spectrum subtends an angle of 42.25^o, the violet end an angle of 40.58^o, giving a bow of width c. 1.7^o. This is the *primary rainbow.* [**1**]

Sometimes one can see a *secondary rainbow*, with *two* reflections rather than one: larger, fainter, and with the order of the colours *reversed*. [**1**]

The first qualitatively correct explanation is due to *M. A. de Dominis* (1564-1624) in 1611 (there is also work by Kepler in 1611). [**1**]

This was taken further by Snel in 1618 and 1621. [**1**]
René Descartes (1596-1650); *Discours de la Méthode ...*, 1637; (first appendix, *La Géométrie*, cartesian geometry;) second appendix, *La Dioptrique*. This gave a treatment of the rainbow, including an estimate of the angle, 42^o. Descartes did not have calculus, so could not give an analytic solution. [**1**]
(Sir) Isaac Newton (1642-1727). Newton lectured in Cambridge (1669-71) on the rainbow, and wrote a paper for the Royal Society in 1672 on the composite nature of white light. His *Principia* of 1687 contains all this. [**1**]

Source: C. B. BOYER, *The rainbow: From myth to mathematics*, Princeton UP, 1987 (1st ed. 1959).
[Mentioned in lectures, though less detail was given there.]

Q6. *The formula of spherical excess.*
Albert Girard (1590-1633) of Flanders (B 16.4)
Invention nouvelle en l'algèbre (1629): Negative and imaginary roots of polynomials; irreducible case of the cubic; sums of roots, of squares of roots, etc. Girard also conjectured the Fundamental Theorem of Algebra.

The book also contains *Girard's formula*, or the *formula of spherical excess.* For a spherical triangle with angles A, B, C, $A + B + C > \pi$, and $A + B + C - \pi$ is called the *spherical excess.* If the sphere has radius r, then the area Δ of the triangle is given by

$$\Delta = r^2(A + B + C - \pi).$$

Note: So for small spherical triangles on the earth's surface, the sum of the angles is approximately π. This just says that we may neglect the Earth's curvature for triangles small in relation to the Earth, and this is of course used in map making (recall the trig points on OS maps!). **[4]**

On a sphere, a *lune* is the region between two great circles. The ratio of the area of the "A-lune" to that of the sphere is A/π (draw a diagram), and similarly for the B- and C-lunes. If we sum the areas of the three lunes, we cover the area of the sphere, but that of the spherical triangle ABC and its antipodal triangle three times (draw a diagram), giving a sum of $S + 4\Delta$ (where S, Δ are the areas of the sphere and triangle). Divide by $S = 4\pi r^2$:

$$\frac{A}{\pi} + \frac{B}{\pi} + \frac{C}{\pi} = 1 + \frac{4\Delta}{4\pi r^2} : \qquad \Delta = r^2(A + B + C - \pi). \qquad \textbf{[3]}$$

C. F. Gauss (1777-1855) and *Differential Geometry.*
Disquisitiones generales circa superficies curvas (1827).

This covers parametric representation of surfaces; curvilinear coordinates; conformal property; Gaussian curvature K; the Gauss-Bonnet theorem (extended by Pierre Bonnet (1819-1892)in 1848). For geodesic triangles with angles A, B, C, $A + B + C - \pi = \int K dS$, with dS the 'element of surface area'. For a sphere (constant positive curvature), one recovers Girard's formula of spherical excess. The plane has zero curvature, and $A + B + C = \pi$ (Euclid, Book I Prop. 32). The case $K < 0$ involves *non-Euclidean geometry*, a possibility that Gauss hinted at in a letter of 1799 to Wolfgang Bolyai, father of Johan (Janos) Bolyai (co-discoverer of non-Euclidean geometry, with Lobachevski). **[3]**
[Seen – lectures]

Q7. *The emergence of the decimal system.*
Aryabhata, author of Aryabhatiya (499 AD) **[2]**

Here we find decimal-place notation: 'from place to place each is ten times the preceding'. The modern decimal numerals 1,2,3,4,5,6,7,8,9 are loosely called Arabic in English, but are called Hindu in Arabic; perhaps 'Hindu-Arabic' would be better. These evolved gradually; the key recognition that by use of place notation the same symbol could be used for three as for thirty, etc., had taken place by 595 AD (Indian source: date 346 in decimal notation), and in Western sources by 662 (Sebokt of Syria).

The zero symbol 0 came later. It had emerged by 876 in India (on an inscription in Gwalior), with the modern 0 for zero. The key components of (i) decimal base, (ii) positional notation, (iii) symbols for 0,1,2,...,9 were thus all in place. It seems that the Hindus did not invent any of them, but they did integrate them into (essentially) their modern form.

Fibonacci, Leonardo of Pisa (c.1180-1250) **[2]**

Leonardo of Pisa, son of Bonaccio (hence 'Fibonacci') wrote the *Liber Abaci* (Book of the Abacus) in 1202. This was the most influential European mathematical work before the Renaissance, and was the first such book to stress the value of the (Hindu-)Arabic numerals (Fibonacci had studied in the Muslim world and travelled widely in it).

Nicholas Chuquet (fl. c. 1500) **[2]**
Triparty en la science des nombres, 1484: the most important European mathematical text since the *Liber Abaci*.
Part I: Hindu-Arabic numerals; addition, subtraction; multiplication; division; Part II: Surds; Part III: Algebra; laws of exponents; solution of equations.

Simon Stevin (1548-1620) of Bruges (Brugge), Flemish military engineer. **[2]**
Die Thiende, *La disme* ('the tenth'), 1585.

Stevin's elementary book did more than any other to popularise the use of decimal fractions (decimals), and to spread awareness of their computational value and superiority.

Napoleon Bonaparte (1769-1821). **[2]**

Napoleon's main contribution to science was that his conquests spread the new metric system, in universal use today.
[Seen – lectures]

Q8. *The mathematics of heat.*
Joseph Fourier (1768-1830); Fourier became a Prefect under Napoleon.
Théorie Analytique de la Chaleur (1822) (*The analytical theory of heat*).

The propagation of heat in a medium is governed by the *heat equation*: with k the thermal diffusivity, the temperature u satisfies

$$\nabla^2 V := \partial^2 V/\partial x^2 + \partial^2 V/\partial y^2 + \partial^2 V/\partial z^2 = k^{-1}\partial V/\partial t.$$

This is the prototypical *parabolic* partial differential equation (PDE). **[3]**

The theory of *Fourier series* – representing a function $f(\theta)$ by a trigonometric series $\sum(a_n \cos n\theta + b_n \sin n\theta)$ – is closely connected with the study of heat propagation, and both were developed by Fourier, in his pioneering book and earlier. The continuous analogue of this is the *Fourier integral*,

$$\hat{f}(t) := \int_{-\infty}^{\infty} e^{ixt} f(x) dx.$$

Applying the Fourier transform can transform a PDE into an ordinary differential equation (ODE) – much easier to solve. Example: solution of the heat equation. **[2]**

Just as heat *diffuses* through a conducting medium, so the mathematics of *diffusion* is dominated by the Fourier transform. Example: *diffusions* in Probability Theory – prototype: Brownian motion. This originates in the work of *Robert Brown* in 1828 (observing the movement of pollen particles suspended in water); the mathematical theory of Brownian motion as a stochastic process is due to Norbert Wiener (1894-1964) in 1923. **[2]**
Rudolf Clausius (1822-1888) *and Thermodynamics*
Über die bewegende Kraft der Wärme (1850) [On the moving force of heat]
This closes with the most famous two-sentence passage in the history of science: Die Energie der Welt ist konstant.
Die Entropie der Welt strebt einem Maximum zu.
[The energy of the world is constant. The entropy of the world strives towards a maximum. (Read 'universe' for 'world' in each here.)]

These are the First and Second Laws of Thermodynamics. The first is the Law of Conservation of Energy. The second is that *entropy increases*. Entropy is a measure of disorder: nature 'moves towards disorder', as the vast majority of possible states are disordered. Thermodynamics is at the heart of heat transfer, which is a large part of Chemical Engineering.
[Seen – lectures] **[3]**

Q9. *The Prime Number Theorem (PNT).*

PNT states that

$$\pi(x) := \sum\nolimits_{p \le x} 1 \sim li(x) := \int_2^x dt/\log t \sim x/\log x \qquad (x \to \infty) \quad (PNT)$$

This was conjectured on numerical grounds by GAUSS (c. 1799; letter of 1848) and A. M. LEGENDRE (1752-1833; in 1798, *Essai sur la Théorie des Nombres*). **[2]**

PNT was proved independently in 1896 by J. HADAMARD (1865-1963, French) and Ch. de la Vallée Poussin (1866-1962, Belgian). Both used Complex Analysis and ζ. **[2]**

In 1737 L. EULER (1707-1783) found his Euler product, linking the primes to $\sum_{n=1}^{\infty} 1/n^{\sigma}$ for real σ (later the Riemann zeta function). **[1]**

In 1859 B. RIEMANN (1626-66) studied

$$\zeta(s) := \sum\nolimits_{n=1}^{\infty} 1/n^s \qquad (s \in \mathbb{C})$$

using Complex Analysis (M2PM3), then still fairly new, developed by A. L. CAUCHY (1789-1857), 1825-29. He showed the critical relevance of the *zeros of* $\zeta(s)$ to the *distribution of primes*. One can show that:
(i) ζ can be continued analytically from *Re* $s > 1$ to the whole complex plane $\mathbb{C}$, where it is holomorphic except for a simple pole at 1 of residue 1 (III.3);
(ii) The only zeros of ζ outside the *critical strip* $0 < \sigma = Re\ s < 1$ are the so-called *trivial zeros* $-2, -4, \dots, -2n, \dots$ (trivial in that they follow from the *functional equation* for ζ;
(iii) PNT is closely linked to non-vanishing of ζ on the 1-line: $\zeta(1 + it) \neq 0$.

The *Riemann Hypothesis* (RH) of 1859 is that the only zeros of ζ in the critical strip are on the *critical line* $\sigma = \frac{1}{2}$. RH is still open, and is the most famous and important open question in Mathematics. Its resolution would have vast consequences for prime-number theory (especially error terms in PNT). **[4]**

Since counting primes relates to $\mathbb{N}$ ($\subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$), it seemed strange and unaesthetic to use complex methods. Great efforts were make to provide an *elementary proof*, over half a century.This was done in 1948 by A. SELBERG (1917-2009) and P. ERDÖS (1913-1996) (published independently in 1949). **[1]**

[Seen in lectures, in less detail. The above is taken from my Analytic Number Theory notes, Lecture 15.]

Q10. *The mathematics of length, area and volume.*
Greek geometry.

Geometry was one of the strengths, and main interests, of (ancient) Greek mathematics. They knew a great deal (several of Euclid's books are on geometry), including many highly non-trivial area and volume calculations. Thus they knew integral calculus, without the name (and without differential calculus) – Archimedes' method of exhaustion, etc.

Greek definitions of geometrical entities were sketchy: a point is that with no extent; a line is that with no thickness, etc. **[3]**

Henri Lebesgue (1875-1941)

Lebesgue took his PhD in 1902, supervised by Emile Borel (1871-1956). In his pioneering thesis 'Intégrale, longueur, aire' (*Annali di Mat.* **7** (1902), 231-259), Lebesgue introduces the new subject of *measure theory*, and a new integral, now called the *Lebesgue measure*. Despite being harder to set up than the Riemann integral, it is much more general and powerful and much easier to manipulate (e.g., in interchanging limit and integral – *Lebesgue's monotone convergence theorem* and *Lebesgue's dominated convergence theorem*). Lebesgue measure is the mathematics of length, area and volume. It is also the mathematics of gravitational mass, electrostatic charge, and probability; in probability, the total mass is one, and the integral is the expectation. The essence of measure theory is *countable additivity*: for A_n disjoint measurable sets with measures $\mu(A_n)$, the measure of their union is the sum of their measures:

$$\mu(\bigcup_0^\infty A_n) = \sum_0^\infty \mu(A_n). \qquad [\mathbf{5}]$$

Non-measurable sets.

With the development of modern set theory, beginning with Cantor, foundational questions such as the Zermelo-Fraenkel axioms (ZF), the Axiom of Choice (AC), and their union (ZFC) (the ordinary assumptions of modern mathematics) were studied. It was shown by Vitali that, assuming (AC), there exist *non-measurable sets*. Thus some subsets of the plane are too irregular to have an area, etc. Indeed, there is a sense in which (under (AC)) *most* sets are non-measurable. On the other hand, it is consistent to assume axioms other than (AC), under which *all* sets are measurable. **[2]**

[Seen – lectures.]

Section B.

Q1 *Conics and orbits: Kepler and Newton.*
(i) *Polar equation of a conic.*

With focus F, directrix L distant ℓ (the latus rectum) from P, the focus-directrix property $PF = e.PL$ (Pappus, c. 290 AD) is $r = \ell - er\cos\theta$:

$$r(1 + e\cos\theta) = \ell, \qquad \frac{1}{r} = \frac{1}{\ell}(1 + e\cos\theta). \qquad [\mathbf{4}]$$

(ii) We need the components of velocity and acceleration along and perpendicular to the radius vector OP in polar coordinates:

$$x = r\cos\theta, \qquad y = r\sin\theta;$$

$$\dot{x} = \dot{r}\cos\theta - r\sin\theta\dot{\theta}, \qquad \dot{y} = \dot{r}\sin\theta + r\cos\theta\dot{\theta};$$

$$\ddot{x} = \ddot{r}\cos\theta - 2\dot{r}\dot{\theta}\sin\theta - r\cos\theta(\dot{\theta})^2 - r\sin\theta\ddot{\theta},$$

$$\ddot{y} = \ddot{r}\sin\theta + 2\dot{r}\dot{\theta}\cos\theta - r\sin\theta(\dot{\theta})^2 + r\cos\theta\ddot{\theta}.$$

So the components of velocity along and perpendicular to OP are $\dot{x}\cos\theta + \dot{y}\sin\theta = \dot{r}$ and $-\dot{x}\sin\theta + \dot{y}\cos\theta = r\dot{\theta}$ (both obvious). The components of acceleration are:

$$\ddot{x}\cos\theta + \ddot{y}\sin\theta = \ddot{r} - r(\dot{\theta})^2 \quad \text{along } OP, \qquad [\mathbf{2}]$$

$$-\ddot{x}\sin\theta + \ddot{y}\cos\theta = 2\dot{r}\dot{\theta} + r\ddot{\theta} = \frac{1}{r}\frac{d}{dt}(r^2\dot{\theta}) \quad \text{perpendicular to } OP \qquad [\mathbf{2}]$$

(no $\ddot{r}$ – 1st – term; 2nd term $2\dot{r}\dot{\theta}$ $(\cos^2 + \sin^2 = 1)$; no $r(\dot{\theta})^2$ – 3rd – term; 4th term $r\ddot{\theta}$ $(\cos^2 + \sin^2 = 1)$).

(iii) If the force is *central* (along OP), then there is no acceleration perpendicular to OP. So $h := r^2\dot{\theta}$ is constant (the angular momentum per unit mass). But if A is the area swept out by the radius vector, $dA = \frac{1}{2}r.rd\theta$: $\dot{A} = \frac{1}{2}r^2\dot{\theta}$. So

$$\dot{A} = \frac{1}{2}h = constant:$$

the radius vector sweeps out equal areas in equal times (Kepler's Second Law – equivalent to central forces). [**4**]

For a central force: write $u := 1/r$. So by (ii), $\dot{\theta} = h/r^2 = hu^2$,

$$\frac{dr}{dt} = \frac{d}{dt}(1/u) = \frac{d}{du}(1/u)\frac{du}{dt} = -\frac{1}{u^2}.\frac{du}{d\theta}.\frac{d\theta}{dt} = -\frac{1}{u^2}.\frac{du}{d\theta}.hu^2: \quad \dot{r} = -hdu/d\theta.$$

$$\frac{d^2r}{dt^2} = \frac{d}{dt}(-h\frac{du}{d\theta}) = -h\frac{d}{d\theta}(\frac{du}{d\theta}).\frac{d\theta}{dt} = -h^2u^2\frac{d^2u}{d\theta^2}: \quad \ddot{r} = -h^2u^2d^2u/d\theta^2.$$

So by (ii), with a the acceleration a along OP towards O,

$$-a = \ddot{r} - r(\dot{\theta})^2 = -h^2u^2\frac{d^2u}{d\theta^2} - \frac{1}{u}.h^2u^4 : \qquad \frac{d^2u}{d\theta^2} + u = \frac{a}{h^2u^2}. \qquad (*)$$

Newton's Inverse Square Law of Gravity is that the acceleration $a = c/r^2 = cu^2$ for c constant: $c = a/u^2$ is constant. So by $(*)$, this is

$$\frac{d^2u}{d\theta^2} + u = \frac{c}{h^2} = b, \qquad (DE) \quad \mathbf{[4]}$$

say.

(iv) The differential equation (DE) has general solution $u = b + c_1\cos\theta + c_2\sin\theta$. We may choose the initial line $\theta = 0$ to make $du/d\theta = 0$ there; then $c_2 = 0$, and $u = b + c_1\cos\theta$:

$$1/r = b + c\cos\theta.$$

But the polar equation of a conic of eccentricity e and latus rectum ℓ is

$$r(1 + e\cos\theta) = \ell : \qquad \frac{1}{r} = \frac{1}{\ell}(1 + e\cos\theta).$$

This identifies the path of a particle moving under the inverse square law as a *conic*. // **[4]**

(v) This is the most important single result in Newton's *Principia* (1687), itself the most important book in the history of science. Our orbit round the Sun is closed, so being a conic, it is an *ellipse*. The eccentricity of the ellipse gives us our seasons.

Johannes Kepler (1571-1630): Assistant to Tycho Brahe (1546-1601) at Prague Observatory from 1600; succeeded him in 1601; *Astronomia Nova*, 1609, Kepler's first and second laws (*Harmonices mundi*, 1619, Kepler's third law). Brahe spent twenty years observing planetary orbits; Kepler spent twenty years analysing his data.

Newton's derivation was geometrical. He avoided calculus (his 'method of fluxions'), preferring to solve an old problem by established methods.

The geometrical content, as noted, is Pappus' focus-directrix property of conics (Alexandria, c. 290 AD), used via polar coordinates (cartesian geometry, 17th C.).

The differential equations method above (technically easier for a modern audience) belongs to the 18th C. (the Bernoullis, Euler etc.). **[5]**

Seen – lectures + problem sheets.

Q2 *Normal distribution; Central Limit Theorem; Least Squares.*
Abraham de Moivre (1667-1754)
The Doctrine of Chances (DC; 1718; 2nd ed. 1738; 3rd ed. 1756, posth.)

De Moivre was born in France, a Huguenot (= Protestant). He fled to England to escape religious persecution after the revocation of the Edict of Nantes in 1685 by Louis XIV. He was elected FRS in 1697.

In 1733 de Moivre derived the normal curve ("error curve": the term "normal" came later)

$$\phi(x) := e^{-\frac{1}{2}x^2}/\sqrt{2\pi}$$

in a 7-page note in Latin *Approximatio ...* (circulated, but not published). An English translation of the *Approximatio* was incorporated in the 2nd and 3rd editions of DC. De Moivre showed that in n Bernoulli trials with parameter p

$$P((S_n/n - p)\sqrt{n}/\sqrt{pq} \in [a, b]) \to \int_a^b \phi(x)dx \qquad (n \to \infty)$$

(recall that $P(S_n = k) = \binom{n}{k}p^k q^{n-k}$, so the LHS is this summed over k with $np+a\sqrt{npq} \leq k \leq np+b\sqrt{npq}$). This is the special case for Bernoulli trials of the "Law of Errors", or *Central Limit Theorem* (CLT). Note that this refines Bernoulli's theorem (*Ars conjectandi*, 1713) by telling us *how fast* S_n/n tends to p: at rate $1/\sqrt{n}$, with a normal limit: in modern terminology,

$$((S_n/n) - p)\sqrt{n}/\sqrt{pq} \sim N(0, 1). \qquad [\mathbf{6}]$$

One can see that this result will need an estimate for the large factorials occurring in $\binom{n}{k} = n!/(k!(n-k)!)$, as in Stirling's formula, 1730:

$$n! \sim \sqrt{2\pi}e^{-n}n^{n+\frac{1}{2}} \qquad (n \to \infty). \qquad [\mathbf{1}]$$

Adrien-Marie Legendre (1725-1833), Professor at the Ecole Militaire
Nouvelles méthodes pour la détermination des orbites des comètes (1805) (supplement 1806, 2nd suppl. 1820).

The orbits of planets and comets, being ellipses and so conics, are determined by two parameters. For flexibility in choice of frame of reference, it may be useful to use p parameters, $\theta_1, \ldots, \theta_p$, where p is small. With exact measurements, these could be determined by p observations, $y_1, \ldots, y_p$ say, but in practice the y_i are polluted by measurement error. However, we may take n readings where n is large (the larger the better – much larger than p). The usual set-up is

$$y_i = \sum_{j=i}^{p} x_{ij}\theta_j + \epsilon_i \qquad (i = 1, \ldots, n),$$

where the x_{ij} are known, the ϵ_i are independent errors of measurement, and the θ_j are to be estimated. How should one estimate the θ_j? Legendre suggested the *method of least squares*: minimise the sum of squares

$$SS := \sum_i (y_i - \sum_j x_{ij}\epsilon_j)^2 = \sum (\text{observed - expected})^2.$$

The p conditions $\partial SS/\partial\theta_j = 0$ give p simultaneous linear equations in p unknowns, the *normal equations* NE; the solutions $\hat{\theta}_j$ give the *least-squares estimators* for the parameters. The method of least squares, in Statistics, is of enormous practical importance. **[6]**

Carl Friedrich Gauss (1777-1855)

Gauss published his work, on the orbit of Ceres (1801), Pallas (another planetoid) and least squares in
Theoria motus corporum coelestium in sectionibus conicus solem ambientem (1809) (TM).

Legendre's book of 1805 presented and named the method, but did not link it to Probability Theory. Gauss, however, *linked least squares with the normal law.* He said in TM 'On the other hand, our principle (principium nostrum), which we have made use of since the year 1795, has lately been published by Legendre ...'. No doubt Gauss was telling the literal truth. Understandably, however, Legendre felt slighted that Gauss should thus seek to claim priority for something first published by Legendre, and protested, first in private correspondence, and in 1820 publicly in a Second Supplement to his 1805 book. **[6]**

Pierre-Simon de Laplace (1749-1827), Professor at the Ecole Normale and the Ecole Polytéchnique.

Théorie Analytique des Probabilités (1812) (TAP) (2nd ed. 1814, 3rd 1820). Laplace gave (Ch. IV of TAP) a thorough treatment of the *method of least squares* (due to Legendre and Gauss). He brought together the roles of the normal distribution in the Central Limit Theorem and the Method of Least Squares, thus achieving the *Gauss-Laplace synthesis.* **[6]**

[Seen – lectures.]

Q3. *The development of linear algebra.*
Determinants. **[5]**

Determinants (which in modern language technically belong to Multilinear Algebra) may be traced to the work of Leibniz (1646-1716) of 1693 (unpublished till 1880), *Maclaurin* (1698-1746, *Treatise on Algebra*, 1748, posth.), *Cramer* (1704-1752: *Cramer's rule* for solution of linear equations, 1752) and *Lagrange* (1736-1813; geometric work of 1775).

The subject came of age with the 84-page paper of 1812 by *Cauchy* (1789-1857); the term *determinant* comes from the *Disquisitiones arithmeticae* (1801) by *Gauss* (1777-1855).
Matrices. **[5]**

Logically, matrices precede determinants, but historically the order was reversed.
Arthur Cayley (1821-1895);
J. J. Sylvester (1814-1897).

The term 'matrix' was introduced in 1850 by Sylvester, and the theory was developed by Cayley in a series of papers in the 1850s, particularly 1858. This is the source – with *Hamilton*'s lectures of 1853 on quaternions – of the *Cayley-Hamilton theorem*: a matrix satisfies its own characteristic equation.

Other sources from this time include Sylvester's Law of Inertia (1852) and Law of Nullity (1884), and the introduction of *Hermitian* matrices [Hermite, 1855] and *orthogonal* matrices [Hermite, 1854; Frobenius, 1878].

Hermann Grassmann (1809-1877); *Ausdehnungslehre* (1844) [Theory of Extensions]. **[4]**

Grassmann's work, which was geometrically motivated, was a major source from which *linear algebra* – the theory of v ector spaces, linear transformations between them, and the matrices representing these, etc. – developed. Grassmann's work is also the source of *Grassmann algebras* – important in modern multilinear and tensor algebra.

Another source was the work of the English geometer *W. K. Clifford* (1845-1879).
Giuseppe Peano (1858-1932), *Calcolo geometrico secondo l'Ausdehnungslehre de Grassmann, preceduto dalle operazioni della logica deduttiva*, Torino, 1888. **[2]**

Here Peano, explicitly building on Grassmann's work, gives the axiomatic definition of a vector space (whether of finite dimension or not) over the reals, and of linear transformations between them, all in modern notation.

Quaternions. **[2]**

Sir William (Rowan) Hamilton (1806-1865) discovered *quaternions* in 1843, and published his *Lectures on Quaternions* in 1853. Quaternions became widely used, especially in applied mathematics, e.g. Thomson and Tait, *Treatise on Natural Philosphy*, 1867.

Vectors. **[2]**

The American physicist *Josiah Willard Gibbs* (1839-1903) (*Elements of Vector Analysis*, 1881) saw that the sub-theory of quaternions giving ordinary 3-dimensional *vectors* is adequate for most purposes in applied mathematics. The standard machinery of *vector algebra* (vector products, scalar products, triple vector products etc.) and *vector calculus* (grad, div, curl, Δ^2; theorems of Stokes, Gauss and Green) developed so successfully as to have generally superceded quaternions by the end of the 19th C.

The 20th C. **[5]**

Linear Algebra now forms one wing, with Abstract Algebra (groups, rings, fields etc.) of *Modern Algebra*, a core part of the undergraduate curriculum. This subject took shape with

B. L. van der Waerden, *Moderne Algebra*, Vol. I (1930), II (1931).

Van der Waerden (1903-1996, PhD 1925, H. de Vries) was a Dutch mathematician much influenced by Emmy Noether (1882-1935, PhD 1907, P. Gordan). Noether and van der Waerden are the main originators of modern algebra, together with their Göttingen colleague David Hilbert (1862-1943).

Van der Waerden's book was in German, and used Fraktur for the equations, as did its eventual English translation. For such reasons, and through passage of time, it has been supplanted as a text by later books, e.g.

Garrett Birkhoff and Saunders MacLane, *A survey of modern algebra*, 1941/1953;

P. M. Cohn, *Algebra*, Vol. 1 (1974/82), 2 (1977/89), 3 (1991).

In the computer age, Numerical Linear Algebra has become a subject in its own right. Classics include Wilkinson's *The algebraic eigenvalue problem* and Golub & van Loan's *Matrix computation*.

Multivariate Analysis in Statistics makes heavy use of Linear Algebra and Matrix Theory. In particular, *singular values decomposition (SVD)*, not found in Algebra texts a generation ago, is now widely used in Statistics (approximation by rank-one matrices) and Numerical Analysis (because of its good numerical stability).

Q4. *The development of the real number system* $\mathbb{R}$.
The Greeks.

Beginning with *Pythagoras* (c. 580 – c.500 BC) and his school at Croton, the Greek mathematicians dealt with:
(a) *numbers*: first natural numbers $\mathbb{N}$, then integers $\mathbb{Z}$ and rationals $\mathbb{Q}$ by the arithmetic operations;
(b) *lengths*: of geometric line-segments.

No tension between these two was noticed at first, but – perhaps by *Hippasus* (c. 400 BC) – it was observed that such natural *geometric* entities as the length $\sqrt{2}$ of the diagonal of the unit square are *irrational*, and so outside the domain of the number system $\mathbb{Q}$. The same proof [by contradiction] shows that $\sqrt{p}$ is irrational for all primes p. **[2]**

Eudoxus (of Cnidus, c. 408 – c. 335 BC) introduced his *theory of proportion*: "$a/b = c/d$ iff $\forall\ m, n \in \mathbb{N},\ ma \leq / = / \geq nb \ \Rightarrow mc \leq / = / \geq nd$". This enabled a rigorous treatment of proportionality of numbers, or of similarity of geometric figures, to be given. **[2]**

Euclid of Alexandria wrote his *Elements*, Books I-XIII, c. 300 BC. The ordering of the material [esp. Books V, Theory of proportions, and X, Incommensurability and surds] was much affected by the technical problems of being unable to treat incommensurables [or irrationals] rigorously.

The Greeks had problems, because they
(a) did not have an adequate theory of irrationals, but
(b) demanded rigorous proofs. **[2]**

The Middle ages.
Thomas Bradwardine (c. 1290-1349), Archbishop of Canterbury. **[1]**
Tractatus de continuo – regarding the real line as a continuum.
Nicole Oresme (1323-1382). **[1]**
De proportionibus proportionum: contains laws of exponents, and suggesting *irrational* exponents. The book also contains (Ch. III, Prop. V) a realisation that a 'typical' real number is *irrational*.

Calculus and analysis: Newton and Euler. **[5]**
Issac Newton (1642-1727), *De analysi* (MS 1669, publ. 1711); *Leonhard Euler* (1707-1783), *Introductio ad analysin infinitorum*, 1748.

Newton systematically handled infinite processes (power series expansions, calculus) avoided by the Greeks. Euler's *Introductio* took this rather loose manipulation with infinite processes further. Various people raised questions about rigour – e.g. *Bishop Berkeley* (1685-1753: *The Analyst*, 1734 – infinitesimals are "the ghosts of departed quantities"), and *Bernhard Bolzano* (1781-1848: *Paradoxien des Unendlichen*, 1851, posth.).

The construction of the reals $\mathbb{R}$ *in 1872.*
Richard Dedekind (1831-1916), *Stetigkeit und die Irrationalzahlen.* **[4]**

Dedekind's construction of $\mathbb{R}$ centres on the idea of a *Dedekind cut* or *section* (Schnitt). The irrational $\sqrt{2}$ divides the rationals $\mathbb{Q}$ into two classes, those $< \sqrt{2}$ and those $> \sqrt{2}$, either of which serves to identify $\sqrt{2}$. The arithmetic operations can be defined on such cuts, turning the class of cuts into a *field*.
Georg Cantor (1845-1918). **[4]**
Cantor's construction of the reals (*Math. Annalen* **4** (1972), 123-132 and **21** (1883), 545-591).

Call a sequence $a = (a_n)$ of rationals *fundamental*, or *Cauchy*, if $a_{m+n} - a_n \to 0$ $(m, n \to \infty)$, i.e.

$$\forall \epsilon > 0 \quad \exists N = N(\epsilon) \ s.t. \ n > N, \ m \geq 0 \ \Rightarrow \ |a_{m+n} - a_n| < \epsilon.$$

Call $a = (a_n)$, $b = (b_n)$ *equivalent* if $a - b := (a_n - b_n) \to 0$ (this *is* an equivalence relation – check). Cantor *defines* a real number to be an *equivalence class of Cauchy sequences of rationals.*

Dedekind's construction is specific to $\mathbb{R}$, as it depends on the *total ordering* of the reals. Cantor's construction is by *completion*, and extends to any *metric space*. We quote that any *complete ordered field* is algebraically isomorphic to $\mathbb{R}$, and so may be identified with it. **[2]**

The 20th Century. **[2]**

Modern Mathematical Logic and Model Theory threw new light on $\mathbb{R}$ – e.g., through the independence of the Axiom of Choice on Zermelo-Fraenkel set theory (Paul J. Cohen (1934-2007) in 1963). See e.g.
T. BARTOSZYNSKI and H. JUDAH, *On the structure of the real line*, Taylor & Francis, 1995.
[Seen – lectures] N. H. Bingham